The Effect of DNA Sequence Divergence on Sexual Isolation in Bacillus

Michael S. Roberts and Frederick M. Cohan

Department of Biology, Wesleyan University, Middletown, Connecticut 06459–0170

Manuscript received August 4, 1992
Accepted for publication February 13, 1993

ABSTRACT

We have investigated the relationship between sexual isolation and DNA sequence divergence in the transformation (at locus *rpoB*) of a naturally competent strain of *Bacillus subtilis*. Using both genomic DNA and a PCR-amplified segment of gene *rpoB* as donor, we found that the extent of sexual isolation at locus *rpoB* was closely predicted, over three orders of magnitude, as a log-linear function of sequence divergence at that locus. Because sexual isolation between a recipient and any potential donor may be determined as a general mathematical function of sequence divergence, transformation is perhaps the only sexual system, in either the prokaryotic or the eukaryotic world, in which sexual isolation can be predicted for a pair of species without having to perform the cross. These observations suggest the possibility of a general approach to the indirect prediction of sexual isolation in bacteria recombining principally by natural transformation.

PROKARYOTIC transformation is unique among all known systems of natural recombination in the simplicity of barriers to genetic exchange: in transformation, sexual isolation between populations sharing the same microhabitat is caused only by differences between recipient and donor in the DNA itself. This is because in transformation it is naked DNA alone that is taken up from the environment, processed and integrated into the genome.

In transformation, sexual isolation may result from two classes of differences in DNA. First, differences between recipient and donor in their restriction-modification systems may cause unmodified (or inappropriately modified) donor DNA to be digested by the recipient's restriction endonuclease (WILSON and YOUNG 1972; COHAN, ROBERTS and KING 1991). Nevertheless, differences between recipient and donor in their restriction systems appear to have only a marginal effect on transformation efficiency, reducing heterogamic transformation by only a factor of three (COHAN, ROBERTS and KING 1991) or even less (TRAUTNER *et al.* 1974; HARRIS-WARRICK and LEDERBERG 1978; IKAWA, SHIBATA and ANDO 1980). Second, and more importantly, divergence between recipient and donor in the DNA sequence itself may hinder recombination. In some transformation systems, such as that of *Haemophilus influenzae*, a short recognition sequence is necessary for entry of donor DNA into the recipient cell (*e.g.*, an 11-bp sequence in the case of *H. influenzae*; Danner *et al.* 1980). Even when the donor DNA contains the recognition sequence required for uptake, the overall sequence divergence between a donor segment and the recipient's homolog may severely reduce the rate of transformation (ALBRITTON *et al.* 1984). In the case of *Bacillus subtilis*, sequence divergence between recipient and donor does not affect entry into the cell as no recognition sequence is involved; however, the overall divergence between a donor segment and the recipient's homolog lowers the stability of the donor-recipient heteroduplex (an intermediate step in integration) and thereby lowers the probability of recombination (te Riele and Venema 1982a, 1982b, 1984). In the Bacillus system, sequence divergence is thought to be responsible for severe reductions in transformation rates, especially in transformation between named species (MARMUR, SEAMAN and LEVINE 1963; DUBNAU *et al.* 1965; GOLDBERG, GWINN and THORNE 1966; WILSON and YOUNG 1972; HARFORD and MERGEAY 1973; SEKI, OSHIMA and OSHIMA 1975; te Riele and VENEMA 1982a; RUDNER, TACKNEY, AND GOTTLIEB 1982; COHAN, ROBERTS and KING 1991). For example, HARFORD and MERGEAY (1973) found that donor DNA from *B. licheniformis* transforms *B. subtilis* at the highly conserved gene *rpoB* at a rate 150 times lower than does *B. subtilis* DNA, but at the poorly conserved gene *trpC* transformation is reduced by a factor of more than 250,000.

Because DNA sequence divergence plays an important role in determining the extent of sexual isolation in transformation, transformation may perhaps prove to be the one system in which sexual isolation can be predicted for a pair of species without having to perform the cross. Provided that sexual isolation between a recipient strain and any potential donor can be determined as a general mathematical function of sequence divergence, one may predict the sexual isolation between a pair of species directly from their DNA sequence divergence. To this end, we have investigated the quantitative relationship between sex-

TABLE 1

Sequences compared in *rpoB* fragment (1102-4468)

Enzyme	Position	Strain										
		1A2	RO-A-4	RO-NN-1	2A2	RO-E-2	RO-C-2	RO-H-1	RO-QQ-2	*B.amy.*	*B.atro.*	*B.lich.*
*Dpn*II	1313	–	–	–	–	–	–	–	–	+	–	–
	1418	–	–	–	–	–	–	–	–	–	–	+
	1495	–	–	–	–	–	–	–	–	–	+	–
	1649	+	+	+	+	+	+	+	+	+	+	+
	1706	+	+	+	+	+	+	+	+	–	+	+
	1721	+	+	+	+	+	+	+	+	+	–	+
	1877	+	+	+	+	+	+	+	+	+	+	–
	1892	+	+	+	–	–	+	+	+	+	–	+
	1982	+	+	+	–	–	+	+	+	+	+	+
	2156	+	+	+	+	+	+	+	+	+	–	+
	2185	–	–	–	–	–	–	–	–	–	+	+
	2300	+	+	+	+	+	–	–	–	–	+	–
	2440	+	+	+	+	+	+	+	+	–	–	+
	2510	+	+	+	–	–	–	–	–	+	+	+
	2677	+	+	+	+	+	+	+	+	+	+	+
	2747	–	–	–	–	–	+	+	+	–	–	–
	3118	–	–	–	–	–	–	–	–	+	–	–
	3280	+	+	+	+	+	+	+	+	–	–	+
	3376	–	–	–	–	–	–	–	–	+	–	–
	3614	+	+	+	+	+	+	+	+	+	+	+
	3850	–	–	–	–	–	–	–	–	–	–	+
	3952	–	–	–	–	–	–	–	–	+	–	–
	4207	+	+	+	+	+	–	–	–	+	–	+
	4248	–	–	–	–	–	–	–	–	–	–	+
	4432	–	–	–	–	–	–	–	–	+	–	–
*Hin*fI	1334	+	+	+	+	+	+	+	+	+	+	+
	1714	+	+	+	+	+	+	+	+	+	–	–
	1862	+	+	+	+	+	+	+	+	+	+	+
	1942	+	+	+	+	+	–	–	–	–	+	+
	2020	+	+	+	+	+	+	+	+	+	+	–
	2070	+	+	+	+	+	+	+	+	–	+	–
	2122	+	+	+	+	+	+	+	+	–	–	–
	2404	–	–	–	–	–	–	–	–	–	–	+
	2518	–	–	+	–	–	+	–	+	–	–	–
	2650	–	–	–	–	–	–	–	–	–	–	+
	2707	–	–	–	–	–	–	–	–	–	+	–
	2775	+	+	+	+	+	+	+	+	+	+	–
	2835	–	–	–	–	–	+	+	+	–	–	–
	2993	+	+	+	+	+	+	+	+	+	+	+
	3101	+	+	+	+	+	+	+	+	+	+	–
	3211	–	–	–	–	–	–	–	–	–	–	+
	3463	–	–	–	–	–	–	–	–	–	+	–
	3482	+	+	+	+	+	+	+	+	+	+	–
	3579	+	+	+	–	–	–	–	–	–	+	–
	3750	–	–	–	–	–	–	–	–	–	+	–
	3907	+	+	–	–	–	–	–	–	+	+	–
	3967	–	–	–	+	+	+	+	+	–	–	–
	4182	–	–	–	+	+	–	–	–	+	–	–
	4273	–	–	–	–	–	–	–	–	–	–	+
	4378	–	–	–	–	–	+	+	+	–	–	–
*Msp*I	1325	–	–	–	–	–	–	–	–	–	–	+
	1404	–	–	–	–	–	–	–	–	–	–	+
	1461	–	–	–	–	–	–	–	–	+	–	–
	1667	+	+	+	+	+	+	+	+	+	+	+
	1821	–	–	–	–	–	–	–	–	+	–	–
	1992	–	–	–	–	–	–	–	–	+	–	–
	2371	–	–	–	–	–	–	–	–	+	–	–
	2567	+	+	+	+	+	+	+	+	+	+	+
	2592	+	+	+	+	+	+	+	+	–	+	–
	3001	–	–	–	–	–	–	–	–	–	–	+
	3053	+	+	+	–	–	+	+	+	+	+	–

TABLE 1

Continued

Enzyme	Position	Strain										
		1A2	RO-A-4	RO-NN-1	2A2	RO-E-2	RO-C-2	RO-H-1	RO-QQ-2	*B.amy.*	*B.atro.*	*B.lich.*
	3175	−	−	−	−	−	−	−	−	+	+	−
	3364	−	−	−	−	−	−	−	−	−	−	+
	3574	−	−	−	−	−	−	−	−	−	−	+
	3586	−	−	−	−	−	−	−	−	+	+	−
	3803	−	−	+	−	−	−	−	−	−	−	−
	3873	+	+	+	+	+	+	+	+	+	+	+
	3914	−	−	−	−	−	+	+	+	−	+	−
	4064	−	−	−	−	−	−	−	−	−	−	+
	4116	+	+	+	+	+	+	+	+	−	+	+
	4161	−	−	−	−	−	+	+	+	−	−	−
	4254	−	−	−	−	−	−	−	−	−	−	+
*Rsa*I	1464	+	+	+	+	+	+	+	+	+	+	−
	1511	−	+	+	−	−	−	−	−	−	−	−
	1640	−	−	−	−	−	+	+	+	−	−	+
	1741	−	−	−	+	+	−	−	−	−	−	−
	1912	−	−	−	−	−	−	−	−	−	+	−
	2048	+	+	+	+	+	+	+	+	+	−	−
	2089	−	−	−	−	−	−	−	−	−	−	+
	2267	−	−	−	−	−	−	−	−	+	+	−
	2603	−	−	−	−	−	−	−	−	+	+	−
	3038	−	−	−	−	−	−	−	−	−	−	+
	3072	−	−	−	−	−	−	−	−	−	+	−
	3223	−	−	−	−	−	−	−	−	+	−	−
	3258	+	+	+	+	+	+	+	+	+	+	−
	3731	−	−	−	−	−	−	−	−	−	−	+
	3821	+	+	+	+	+	+	+	+	+	+	−
	3980	+	+	+	+	+	+	+	+	+	+	+
	4158	+	+	+	+	+	+	+	+	+	+	−
	4201	+	+	+	+	+	+	+	+	+	+	+
	4351	−	−	−	−	−	−	+	+	−	−	−

Presence (+) and absence (−) of recognition sites within the 3366-bp PCR-amplified segment of *rpoB*.

ual isolation and DNA sequence divergence in the transformation of a naturally competent strain of *Bacillus subtilis*.

Our approach was to transform laboratory strain 1A2 of *B. subtilis* (derived from strain 168) with donor DNA from other members of the species and with DNA from a series of species with various degrees of sequence divergence from *B. subtilis*. This strain was chosen as the recipient for our study because it has a high level of competence and because it does not possess restriction activity (COHAN, ROBERTS and KING 1991). Thus, sexual isolation between 1A2 and the various donors may be reasonably ascribed to sequence divergence. From each donor strain (including strain 168 itself) we selected for spontaneous mutations conferring resistance to rifampicin, a resistance phenotype coded only by locus *rpoB* (DEAN and ZEIGLER 1989). DNA from the *rpoB* gene of each rifampicin-resistant mutant was amplified by the polymerase chain reaction (PCR) and used to transform strain 168 toward rifampicin resistance. Sexual isolation (ρ) was quantified as the factor by which the rate of heterogamic transformation toward rifampicin resistance was reduced, compared with the rate of homogamic transformation. Sequence divergence (π) between the donor's *rpoB* fragment and the recipient's homolog was estimated by a restriction digest analysis of the amplified segment of *rpoB*. We will demonstrate a log-linear relationship between sequence divergence (π) and sexual isolation (ρ) for Bacillus.

MATERIALS AND METHODS

Strains: *Bacillus subtilis* laboratory strains 1A2 (a derivative of strain 168) and 2A2 (a derivative of strain W23) were obtained from the Bacillus Genetic Stock Center (BGSC). *B. amyloliquefaciens* type strain 23350 and *B. licheniformis* type strain 14580 were obtained from the American Type Culture Collection (ATCC). *B. atrophaeus* type strain NRRL NRS-213 was kindly provided by LAWRENCE NAKAMURA. We included six Bacillus strains isolated from a natural population in the Mojave Desert (as reported by Cohan, Roberts and KING 1991) on the basis of their phylogenetic clustering: two strains from the "168 subgroup" of *B. subtilis* (strains RO-A-4 and RO-NN-1), one strain from the "W23 subgroup" of *B. subtilis* (strain RO-E-2) and three strains from the "ROH1 subgroup" (strains RO-C-2, RO-H-1 and RO-QQ-2). The division of *B. subtilis* into these subgroups is based on a phylogenetic analysis of genes *gyrA*, *polC* and *rpoB* from 118 strains of *B. subtilis* (COHAN, ROBERTS and

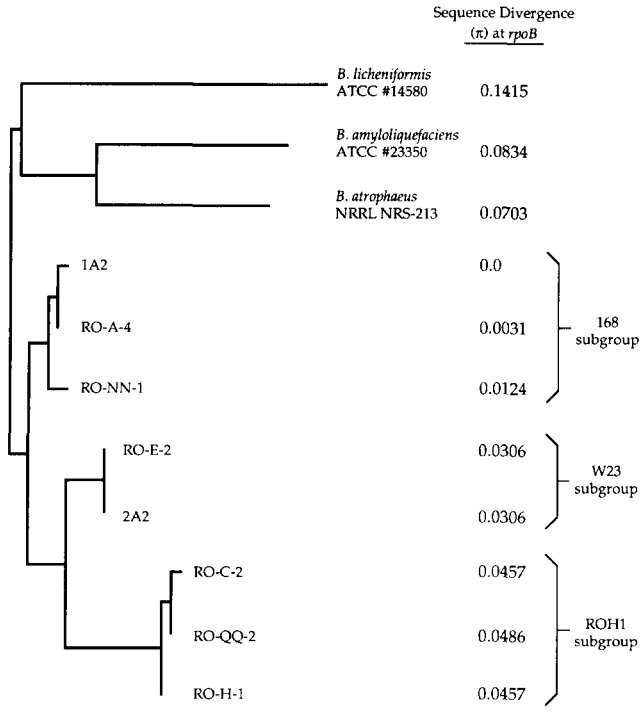

FIGURE 1.—One of two maximum parsimony phylogenies based on the 3366-bp PCR-amplified segment of *rpoB*. The levels of DNA sequence divergence (π) at *rpoB* between each strain and strain 1A2 are estimated from the numbers of shared and nonshared restriction recognition sites for the 3366-bp PCR-amplified segment of *rpoB*. [It should be noted that phylogenies based on PCR-amplified segments of *gyrA* and *polC* each show the 168 and W23 subgroups to be a monophyletic group distinct from the ROH1 subgroup (data not shown). Based on DNA-DNA hybridization results obtained by L. NAKAMURA (unpublished data), we now consider the ROH1 subgroup to be a separate species.]

KING 1991; M. ROBERTS and F. COHAN, unpublished data).

Isolation of rifampicin-resistant mutants: An overnight culture of each of the above strains was streaked onto tryptose blood agar base (TBAB, Difco) supplemented with 10 μg/ml rifampicin (Sigma Chemical Co.). A single rifampicin-resistant colony was then collected for each strain and streaked for isolation on TBAB (without rifampicin) twice. The rifampicin resistance of each isolate was confirmed by plating on TBAB with rifampicin.

Purification of genomic DNA: Genomic DNA from each rifampicin-resistant strain was extracted, purified and ethanol-precipitated as described by COHAN, ROBERTS and KING (1991). For each DNA preparation, the protein:nucleic acid ratio was assayed optically (as the ratio OD_{260}/OD_{280}), and a sample was separated by electrophoresis on a 1% agarose gel (SeaKem GTG, FMC Corporation) and stained with 0.5 μg/ml ethidium bromide. DNA fragments yielded an OD_{260}/OD_{280} ratio greater than 1.5 and were of uniform length of about 50 kb. Genomic DNA preparations were used both as donor DNA in transformation and as template DNA for amplification by PCR.

PCR-amplification of *rpoB* segments: The *rpoB* gene of each rifampicin-resistant (rif^R) mutant was amplified by PCR using genomic DNA as template. Oligonucleotide primers were designed from the *rpoB* coding sequence of *B. subtilis* strain 168 (sequence kindly provided by KATHRYN BOOR, MARIAN DUNCAN and CHESTER W. PRICE) and synthesized by Oligos, etc. The primers extend from bp 1102–1125 (5′-AGG TCA ACT AGT TCA GTA TGG ACG-3′) for the leading strand primer and from bp 4468–4445 (5′-AAC ACC TGG TTC AGG AAC ATT GTC-3′) for the lagging strand primer. As the amino-acid-coding region of *rpoB* extends from bp 1097–4675, the amplified fragment represents 94% of the gene (*i.e.*, 3367 of the 3579 amino-acid-coding nucleotides). PCR was performed according to the GeneAmp™ protocol (Perkin-Elmer-Cetus) in 100-μl volume reactions of PCR buffer (50 mM KCl, 10 mM Tris-HCl, 2 mM $MgCl_2$ and 0.1 mg/ml Bactogelatin) with the following final concentration of reagents: 800 μM dNTPs, 1 mM each primer, 2.5 Units AmpliTaq™ DNA polymerase (Cetus) and ~0.2 μg genomic DNA. Each amplification trial was performed under the same PCR profile of 50 cycles: denaturation at 94° for 30 sec, annealing at 52° for 45 sec and extension at 70° for 3.5 min. Each set of PCR trials included a positive control using as template the genomic DNA of the rif^R mutant of 1A2, and a negative control consisting of all constituent reaction components except for template DNA. The products of each PCR trial were visualized on a 1% agarose gel to determine the length and purity of yield. We used in subsequent experiments only those PCR trials where a single band of 3.4 kb was visualized and where no contamination had occurred in the negative control. Following amplification, each PCR-amplified DNA preparation

TABLE 2

Sexual isolation between recipient strain 1A2 and various donor strains

Donor strain	Source of donor DNA: Genomic DNA	Source of donor DNA: PCR-amplified *rpoB* segment
B. subtilis—168 subgroup		
1A2	0 (15)	0 (18)
RO-A-4	0.02 ± 0.04 (5)	0.43 ± 0.50 (11)**
RO-NN-1	0.06 ± 0.17 (5)	0.23 ± 0.07 (7)***
B. subtilis—W23 subgroup		
2A2	0.10 ± 0.07 (5)*[a]	0.37 ± 0.10 (7)***
RO-E-2	0.19 ± 0.10 (5)**	0.08 ± 0.39 (14)
B. subtilis—ROH1 subgroup		
RO-C-2	0.54 ± 0.12 (5)***	0.59 ± 0.25 (7)***
RO-H-1	0.58 ± 0.07 (5)***	0.27 ± 0.38 (11)*
RO-QQ-2	0.44 ± 0.07 (5)***	0.64 ± 0.09 (7)***
B. atrophaeus		
NRRL NRS-213	0.86 ± 0.07 (5)***	1.18 ± 0.55 (5)***
B. amyloliquefaciens		
ATCC 23350	1.39 ± 0.05 (5)***	1.82 ± 0.31 (5)***
B. licheniformis		
ATCC 14580	2.61 ± 0.19 (5)***	3.46 ± 0.63 (4)***

Strain 1A2 was transformed using either genomic DNA or a PCR-amplified segment of *rpoB* from the donor strain. Donor strains are arranged by DNA-sequence clusters. For each replicate trial, in which one culture of recipient 1A2 was transformed by several donors (including 1A2), sexual isolation was estimated as the ratio of the homogamic to the heterogamic transformation frequency ($\log_{10}$-transformed). The means and standard deviations of the sexual isolation estimates are based on the number of experimental trials reported in parentheses. The sexual isolation between 1A2 as donor and 1A2 as recipient is 0 by definition. Asterisks following sexual isolation parameters indicate significance of deviation from 0 in a one-tailed test.

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.005$.

[a] Not significant after adjustment for multiple comparisons with a sequential Bonferroni test.

TABLE 3

Frequency of transformation ($\log_{10}$-transformed) of recipient strain 1A2

Donor strain	Source of donor DNA: Genomic DNA	Source of donor DNA: PCR-amplified *rpoB* segment
B. subtilis—168 subgroup		
1A2	−1.93 ± 0.25 (15)	−2.26 ± 0.33 (18)
RO-A-4	−1.95 ± 0.15 (5)	−2.73 ± 0.38 (11)
RO-NN-1	−2.24 ± 0.10 (5)	−2.45 ± 0.20 (7)
B. subtilis—W23 subgroup		
2A2	−1.74 ± 0.22 (5)	−2.58 ± 0.23 (7)
RO-E-2	−2.11 ± 0.15 (5)	−2.28 ± 0.23 (14)
B. subtilis—ROH1 subgroup		
RO-C-2	−2.47 ± 0.11 (5)	−2.80 ± 0.27 (7)
RO-H-1	−2.50 ± 0.12 (5)	−2.58 ± 0.31 (11)
RO-QQ-2	−2.36 ± 0.19 (5)	−2.85 ± 0.24 (7)
B. atrophaeus NRRL NRS-213	−2.53 ± 0.21 (5)	−3.53 ± 0.21 (5)
B. amyloliquefaciens ATCC 23350	−3.06 ± 0.19 (5)	−4.18 ± 0.20 (5)
B. licheniformis ATCC 14580	−4.28 ± 0.31 (5)	−5.92 ± 0.21 (4)

Strain 1A2 was transformed by DNA from various donor strains using either genomic DNA or a PCR-amplified segment of *rpoB*. Means and standard deviations are based on the number of experimental trials reported in parentheses. Donor strains are arranged by DNA-sequence clusters.

was extracted, purified and ethanol-precipitated as described above.

Restriction-digest analysis of sequence divergence: An aliquot of each 3.4-kb PCR-amplified *rpoB* fragment was subjected to restriction by each of four 4-cutter restriction endonucleases (*Dpn*II, *Hin*fI, *Msp*I and *Rsa*I; New England Biolabs), as described by COHAN, ROBERTS and KING (1991). Restriction digests were assayed by electrophoresis on 3.5% and 5% acrylamide gels and stained with 0.5 μg/ml ethidium bromide. Fragment lengths were estimated using a *Hae*III digest of φX174RF (New England Biolabs) and an appropriate 4-cutter restriction digest of strain 1A2 as lane markers. Where ambiguities existed in determining the exact size and identity of an *rpoB* digest fragment, smaller PCR fragments within the *rpoB* gene were amplified as described above using internal primers and likewise assayed. Nucleotide substitutions from the *rpoB* sequence of strain 168 were inferred for each strain using a computer algorithm described by COHAN, ROBERTS and KING (1991, Appendix 2).

Transformation: Following a protocol of COHAN, ROBERTS and KING (1991) modified from BOYLAN *et al.* (1972), strain 1A2 was induced to become competent for transformation in liquid culture and transformed to rifampicin resistance with either genomic DNA or the PCR-amplified 3.4 kb *rpoB* fragment from each donor strain. Briefly, the recipient strain was first incubated in a minimal liquid medium at 37° until 90 min after the end of logarithmic phase. The culture was then diluted into a stepdown liquid medium and grown for 60 min. Either genomic or PCR-amplified DNA was added at a final concentration of 9 μg/ml DNA. [We have shown this to be a saturating concentration for recipient strain 1A2, for both genomic DNA (COHAN, ROBERTS and KING 1991) and for PCR-amplified DNA (data not shown).] After a 30-min incubation with DNA, DNase was added at a final concentration of 20 μg/ml. Each culture was then diluted, plated onto TBAB and incubated for 2 hr at 37° prior to the addition of a soft agar (1%) overlay containing rifampicin (50 μg/ml).

The following experiment was performed to demonstrate that transformation by preparations of amplified DNA was almost entirely due to PCR-amplified DNA and not due to the genomic template DNA. For each transformation trial, a separate 1A2 culture was transformed with a 100-μl aliquot of an extracted and purified PCR mixture containing all constituent PCR components and genomic DNA as template, but lacking AmpliTaq™ polymerase (so that the only DNA contributing to transformation was the genomic template DNA). The fraction of transformation from PCR mixtures due to genomic DNA ranged from 1/51 to 1/2399.

RESULTS

Sequence divergence between strain 1A2 and donor strains: A 3366 bp region within the *rpoB* gene (representing 94% of the amino acid-coding region) was amplified by PCR from various subgroups of *B. subtilis*, as well as from the type strains of *B. atrophaeus*, *B. amyloliquefaciens* and *B. licheniformis*. Digestion of each strain's *rpoB* fragment with four restriction enzymes (separately) yielded a total of 91 recognition sites, 80 of which were polymorphic among the strains. The distribution of restriction recognition sites for each enzyme is shown in Table 1. Nucleotide substitutions were inferred from restriction site variation using the computer algorithm described by COHAN, ROBERTS and KING (1991, Appendix 2). The set of nucleotide substitutions at *rpoB* was used to construct a phylogeny of strains (Figure 1), using FELSENSTEIN'S (1989) DNA parsimony algorithm. The DNA sequence divergence (π) for the 3366-bp segment of *rpoB* between strain 1A2 and each of the donor strains was estimated from the numbers of shared and nonshared restriction recognition sites, using Equation 5.50 of NEI (1987), and is shown in Figure 1.

Transformation of strain 1A2 with genomic DNA: Sexual isolation (ρ) between a particular test donor and the recipient 1A2 was quantified as the ratio of the frequency at which the 1A2 rif^R mutant transformed 1A2 (homogamic rate) to the frequency at which the test donor transformed 1A2 (heterogamic rate). This ratio was then log-transformed (yielding $\log_{10}\rho$). In each replicate experimental trial for a test donor, one culture of recipient cells was transformed (separately) with DNA from the 1A2 rif^R mutant as well as with DNA from the test donor. Sexual isolation was determined for each replicate trial, and the distribution of log-transformed sexual isolation estimates is presented for each donor in Table 2. The rifampicin-resistance mutant of 1A2 transformed 1A2 at a baseline level of 1.2×10^{-2} ($=10^{-1.93}$; Table 3). Other strains of *B. subtilis* showed only weak sexual isolation from recipient 1A2, with

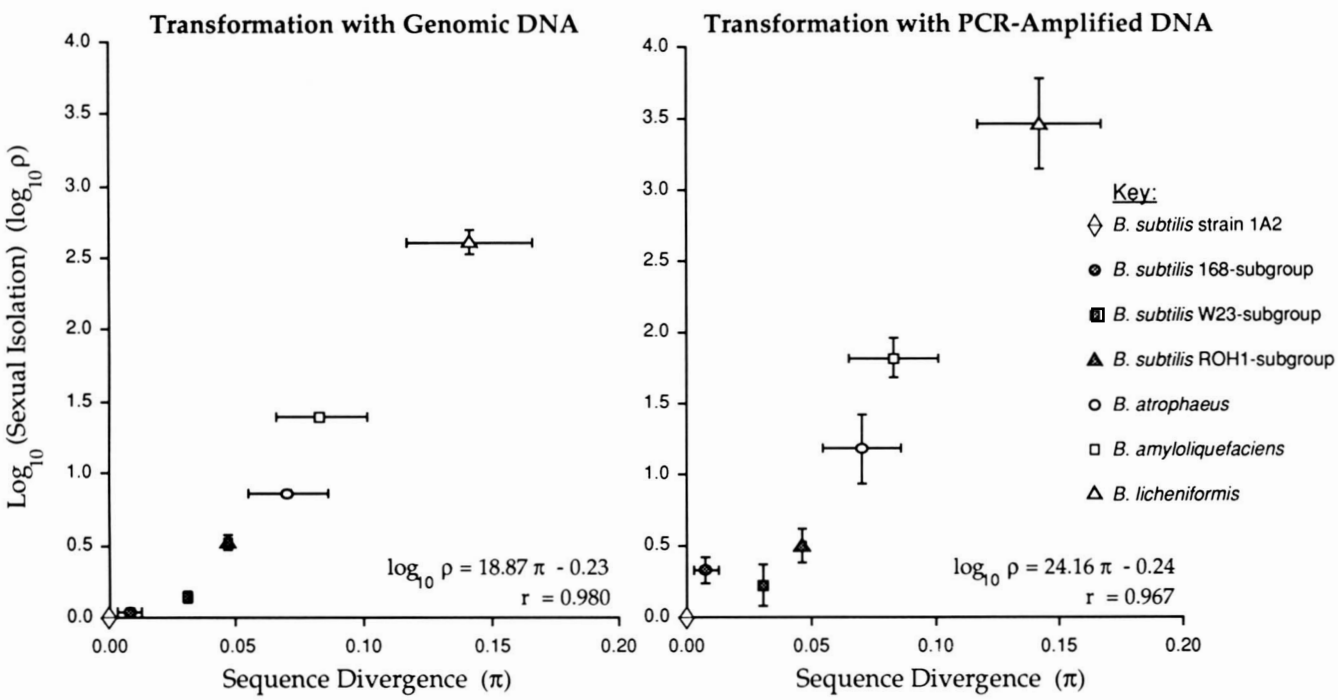

FIGURE 2.—The relationship between log-transformed sexual isolation (log $_{10}\rho$) and DNA sequence divergence (π), with transformation by genomic DNA and by a PCR-amplified segment of *rpoB*. Each sequence cluster of strains is represented by one point. For each sequence cluster in which more than one strain was tested as donor (*i.e.*, 168, W23 and ROH1 subgroups of *B. subtilis*), means and standard errors for sexual isolation are based on the distribution of strain means, and means and standard errors for sequence divergence are based on the distribution of estimates for the various strains. For sequence clusters in which only one strain was tested, means and standard errors for sexual isolation are based on the distribution of estimates from the various experimental trials for that particular strain, and estimates and standard errors for sequence divergence for that strain are based on NEI'S (1987) Equations 5.50, 5.51 and 5.43.

log-transformed sexual isolation estimates ranging from 0.03–0.21 [*i.e.*, reduction of transformation by factors ranging from 1.07 (= $10^{0.03}$]–1.62 [= $10^{0.21}$)]. In only one case (RO-E-2) was the log-transformed estimate of sexual isolation significantly different from zero. Strains from the ROH1 subgroup of *B. subtilis* showed a significant degree of sexual isolation from 1A2 [mean reduction factor of 3.34 (= $10^{0.52}$)], as did all strains from other species [mean reduction ranging from a factor of 7.24 (=$10^{0.86}$) for *B. atrophaeus* to a factor of 407 (=$10^{2.61}$) for *B. licheniformis* (data from Table 2)].

The relationship between sexual isolation and sequence divergence fits the following log-linear regression with a coefficient of determination of 96% ($F_{1,5}$ = 118.3; $P < 0.0001$; Figure 2):

$$\log_{10}\rho = 18.87\pi - 0.23 \qquad (1)$$

in which ρ = sexual isolation (calculated as described above) and π = sequence divergence at *rpoB* (based on restriction digest analysis of the 3366-bp segment).

The negative *y*-intercept of Equation 1 would suggest that the log-linear relationship between isolation and divergence does not ensue until a threshold level of divergence is reached, since a negative level of log-transformed sexual isolation is not expected for low levels of sequence divergence. This threshold hypothesis is not strongly supported, however, because the *y*-intercept (-0.23 ± 0.12) is not significantly different from zero ($P = 0.12$).

Transformation of 1A2 with PCR-amplified fragments of *rpoB*: For each donor strain, the 3366-bp PCR-amplified segment of *rpoB* transformed 1A2 at a lower frequency than did genomic DNA (Table 3). One contributing factor may be that smaller DNA fragments transform at lower efficiency than larger fragments (DUBNAU 1991). In contrast to the case for transformation with genomic DNA, all donors except one (RO-E-2, in the W23 subgroup of *B. subtilis*) are sexually isolated from recipient 1A2 (Table 2).

As for the case of transformation with genomic DNA, the relationship between sexual isolation and sequence divergence at *rpoB* follows a log-linear function with a high (93%) coefficient of determination ($F_{1,5}$ = 71.5; $P = 0.0004$; Figure 2):

$$\log_{10}\rho = 24.16\pi - 0.24. \qquad (2)$$

As was the case for genomic DNA, the negative *y*-intercept weakly suggests that a threshold level of divergence is required before sexual isolation increases in log-linear fashion with increases in sequence divergence, but this hypothesis is not supported by a significant deviation of the *y*-intercept from zero (*y*-intercept: -0.24 ± 0.20; $P = 0.29$).

The relationships between log-transformed sexual isolation and sequence divergence, for genomic DNA and for PCR-amplified DNA, were compared in an analysis of covariance [model: log-transformed sexual isolation depends on the effect of DNA source (class variable), the effect of sequence divergence (continuous variable) and the interaction between DNA source and sequence divergence]. A nonsignificant level of interaction between sequence divergence and source of DNA ($F_{1,10} = 2.51$; $P = 0.14$) suggests that the slope of the relationship between log-transformed sexual isolation and sequence divergence is about the same for the two DNA sources. A marginally significant effect of DNA source ($F_{1,10} = 3.37$; $P < 0.10$) suggests that sexual isolation may have been greater with one source of DNA (PCR-amplified) than with the other (genomic). The effect of sequence divergence was highly significant ($F_{1,10} = 165.6$; $P < 0.0001$), as expected from the high coefficients of determination for Equations 1 and 2.

DISCUSSION

This study has shown that, for the transformation system of *Bacillus subtilis*, the sequence divergence between recipient and donor almost entirely predicts the degree of sexual isolation: with a 93% coefficient of determination, a log-linear function (Equation 2) predicts the extent of sexual isolation at locus *rpoB* from sequence divergence at that locus. This result confirms earlier inferences that sequence divergence is an important determinant of sexual isolation in Bacillus (MARMUR, SEAMAN and LEVINE 1963; DUBNAU *et al.* 1965; GOLDBERG, GWINN and THORNE 1966; WILSON and YOUNG 1972; HARFORD and MERGEAY 1973; HARRIS-WARRICK and LEDERBERG 1978; SEKI, OSHIMA and OSHIMA 1975; TE RIELE and VENEMA 1982a).

The functional relationship found between sequence divergence and sexual isolation is fairly robust with respect to the length of donor DNA fragments, since nearly the same relationship was found for transformation with the PCR-amplified segment of *rpoB* as for transformation with much longer segments of genomic DNA. The slopes of the relationship are about the same for the two sources of DNA, although the extent of sexual isolation is generally greater for PCR-amplified DNA than for genomic DNA [as suggested by a marginally significant ($P < 0.10$) effect of DNA source in the analysis of covariance]. This difference, if real, may be explained by an increased stability of the heteroduplex DNA when anchored by a longer flanking sequence (in the case of genomic DNA).

That a log-linear relationship may be general for other transformation systems is suggested by the only other investigation that has taken a systematic approach to the relationship between the frequency of recombination and sequence divergence. SHEN and HUANG (1986) investigated the effect of sequence divergence on the rate of recombination between genes on a plasmid and genes on a phage in *E. coli*, and found that 10–11% sequence divergence yielded a 55-fold reduction in recombination, 16% divergence yielded a 146-fold reduction, and 30–35% a 300-fold reduction. A log-linear function gives a 77% coefficient of determination, although the regression slope is not significant ($F_{1,2} = 6.79$; $P = 0.12$):

$$\log_{10}\rho = 7.15\pi + 0.53.$$

While the SHEN and HUANG (1986) study does not address the effect of sequence divergence on transformation directly, the study does address the effect of sequence divergence on one component of successful transformation: the probability of recombination once donor DNA has been processed into the cell. Since in the Bacillus system all donor DNA enters the cell with equal probability (DUBNAU 1991), the *E. coli* study of SHEN and HUANG (1986) and the present study are both measuring the extent to which sequence divergence reduces recombination for DNA already in the cell. Although sexual isolation appears to increase with sequence divergence in both systems (although not necessarily in a log-linear fashion for *E. coli*), these systems differ significantly in their sensitivity to sequence divergence (test for difference of slopes in an analysis of covariance: $F_{1,7} = 16.03$; $P = 0.005$). We are currently investigating other bacterial systems with natural transformation to determine whether the log-linear form is universal and, if so, to determine the range of sensitivities of different transformation systems to DNA sequence divergence.

It should not be expected that all transformation systems would be identical in their sensitivity to sequence divergence, as one might expect if the sexual isolation caused by sequence divergence were an intrinsic property of DNA. RAYSSIGUIER, THALER and RADMAN (1989) have shown, for example, that the efficiency of homologous recombination between divergent segments in *E. coli* can be increased by orders of magnitude by a mutation in a proof-reading enzyme. One should expect that differences between transformation systems in the characteristics of their proofreading systems may lead to large differences in their sensitivity to sequence divergence.

The generality of a log-linear effect of sequence divergence on sexual isolation should only be expected to apply to those transformation systems that are like *B. subtilis* in that (1) all DNA enters competent cells with equal efficiency and (2) the reduction in transformation with sequence divergence is due to reduced stability of the heteroduplex. Thus, transformation systems such as that found in Haemophilus

which require a recognition sequence for uptake into the cell (DANNER *et al.* 1980) would not necessarily be expected to follow the log-linear model. [However, for donor DNA segments with the requisite Haemophilus recognition sequence it is possible that sexual isolation may be predicted from sequence divergence by such a model (ALBRITTON *et al.* 1984).] Also, differences in restriction-modification systems may lead to deviations from a general relationship that describes sexual isolation as a log-linear function. This is because strains that are nearly identical in sequence may differ in their restriction-modification systems (COHAN, ROBERTS and KING 1991). Nevertheless, it appears that at least in the case of *Bacillus subtilis*, differences in restriction-modification systems play a very minimal role in sexual isolation (TRAUTNER *et al.* 1974; HARRIS-WARRICK and LEDERBERG 1978; IKAWA, SHIBATA and ANDO 1980; COHAN, ROBERTS and KING 1991).

The high coefficient of determination for the log-linear model in the Bacillus system suggests that one may predict the level of sexual isolation between members of this genus without any knowledge of the interaction between a particular pair of species. This study raises hope that, at least for bacteria that recombine principally by transformation, prediction of sexual isolation in bacteria is a reasonable goal.

We wish to thank ROGER MILKMAN and an anonymous reviewer for their critical consideration of this work and the improvements afforded by their helpful comments. This research was supported by NIH grant GM 39501 and by research grants from Wesleyan University.

LITERATURE CITED

ALBRITTON, W. L., J. K. SETLOW, M. THOMAS, F. SOTTNEK and A. G. STEIGERWALT, 1984 Heterospecific transformation in the genus Haemophilus. Mol. Gen. Genet. **193:** 358–363.

BOYLAN, R. J., N. H. MENDELSON, D. BROOKS and F. E. YOUNG. 1972 Regulation of the bacterial cell wall: analysis of a mutant of *Bacillus subtilis* defective in biosynthesis of teichoic acid. J. Bacteriol. **110:** 281–290.

COHAN, F. M., M. S. ROBERTS and E. C. KING, 1991 The potential for genetic exchange by transformation within a natural population of *Bacillus subtilis*. Evolution **45:** 1393–1421.

DANNER, D. B., R. A. DEICH, K. L. SISCO and H. O. SMITH, 1980 An eleven-base-pair sequence determines the specificity of DNA uptake in Haemophilus transformation. Gene **11:** 311–318.

DEAN, D. H., and D. R. ZEIGLER, 1989 *Bacillus Genetic Stock Center: Strains and Data*, Ed. 4. Ohio State University, Columbus, OH.

DUBNAU, D., 1991 Genetic competence in *Bacillus subtilis*. Microbiol. Rev. **55:** 395–424.

DUBNAU, D., I. SMITH, P. MORELL and J. MARMUR, 1965 Gene conservation in Bacillus species, I. Conserved genetic and nucleic acid base sequence homologies. Proc. Natl. Acad. Sci. USA **54:** 491–498.

FELSENSTEIN, J., 1989 *PHYLIP 3.2 Manual*. University of California Herbarium, Berkeley, CA.

GOLDBERG, I. D., D. D. GWINN and C. B. THORNE, 1966 Interspecies transformation between *Bacillus subtilis* and *Bacillus licheniformis*. Biochem. Biophys. Res. Commun. **23:** 543–548.

HARFORD, N., and M. MERGEAY, 1973 Interspecific transformation of rifampicin resistance in the genus Bacillus. Mol. Gen. Genet. **120:** 151–155.

HARRIS-WARRICK, R. M., and J. LEDERBERG, 1978 Interspecies transformation in Bacillus: sequence heterology as the major barrier. J. Bacteriol. **133:** 1237–1245.

IKAWA, S., T. SHIBATA and T. ANDO, 1980 Genetic studies on site-specific endodeoxyribonucleases in *Bacillus subtilis*: multiple modification and restriction systems in transformants of *Bacillus subtilis* 168. Mol. Gen. Genet. **177:** 359–368.

MARMUR, J., E. SEAMAN and J. LEVINE, 1963 Interspecific transformation in Bacillus. J. Bacteriol. **85:** 461–467.

NEI, M., 1987 *Molecular Evolutionary Genetics*. Columbia University Press, New York.

RAYSSIGUIER, C., D. S. THALER and M. RADMAN, 1989 The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. Nature **342:** 396–401.

RUDNER, R., C. T. TACKNEY and P. GOTTLIEB, 1982 Variations of nucleotide sequences among related Bacillus genomes, pp. 339–351, in *Genetic Exchange: A Celebration and a New Generation*, edited by U. N. STREIPS, S. H. GOODGAL, W. R. GUILD and G. A. WILSON. Marcel Dekker, New York.

SEKI, T., T. OSHIMA and Y. OSHIMA, 1975 Taxonomic study of Bacillus by deoxyribonucleic acid-deoxyribonucleic acid hybridization and interspecific transformation. Int. J. Syst. Bacteriol. **25:** 258–270.

SHEN, P., and H. V. HUANG, 1986 Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. Genetics **112:** 441–457.

TE RIELE, H. P. J., and G. VENEMA, 1982a Molecular fate of heterologous bacterial DNA in competent *Bacillus subtilis*. I. Processing of *B. pumilus* and *B. licheniformis* DNA in *B. subtilis*. Genetics **101:** 179–188.

TE RIELE, H. P. J., and G. VENEMA, 1982b Molecular fate of heterologous bacterial DNA in competent *Bacillus subtilis*. II. Unstable association of heterologous DNA with the recipient chromosome. Genetics **102:** 329–340.

TE RIELE, H. P. J., and G. VENEMA, 1984 Molecular fate of heterologous bacterial DNA in competent *Bacillus subtilis*: further characterization of unstable association between donor and recipient DNA and the involvement of the cellular membrane. Mol. Gen. Genet. **195:** 200–208.

TRAUTNER, T. A., B. PAWLEK, S. BRON and C. ANAGNOSTOPOULOS, 1974 Restriction and modification in *B. subtilis*. Mol. Gen. Genet. **131:** 181–191.

WILSON, G. A., and F. E. YOUNG, 1972 Intergenotic transformation of the *Bacillus subtilis* genospecies. J. Bacteriol. **111:** 705–716.

Communicating editor: A. G. CLARK